**Aquatic Peer Input Workshop**
**EPA Initiative To Revise the**
**Ecological Assessment Process for Pesticides**
**June 22-23, 1999**

**Draft Summary Report**

## INTRODUCTION

On June 22-24, 1999, the Environmental Protection Agency (EPA) sponsored two public workshops as part of its initiative to revise the ecological assessment process for pesticides. The goal of this initiative is to identify, develop, and validate tools and methodologies to conduct probabilistic assessments and improve risk characterization. These assessments would be used to address the impacts of pesticides on nontarget organisms.

A key component of this initiative is the Ecological Committee on FIFRA Risk Assessment Methods (ECOFRAM), which is composed of experts drawn from a variety of stakeholders, including government agencies, academia, environmental groups, industry and others. They are divided into the Aquatic and Terrestrial Workgroups and have been working over the last few months to develop recommendations for EPA's consideration in revising the assessment process. They have discussed potential probabilistic tools and methodologies and have summarized their findings in draft reports.

The purpose of the workshops was to provide peer input on these draft reports. Specifically, the workshops would provide EPA with external scientific review, comment and discussion of the draft reports, which EPA would factor into decisions regarding the implementation of ECOFRAM recommendations. In addition, it would provide the ECOFRAM members with comments to help them finalize the reports.

This report provides a summary of the workshop held to address the Aquatic Draft ECOFRAM report.

## PANEL PRESENTATIONS

### Denise Keehner, U.S. EPA, Environmental Fate and Effects Division, Office of Pesticide Programs

In her opening remarks, Ms. Keehner noted that in May 1996 EPA asked the FIFRA Scientific Advisory Panel to review and comment on the Office of Pesticide Program's ecological risk assessment methods and procedures. While recognizing and generally reaffirming the utility of the assessment process currently used, the Panel encouraged the Agency develop the tools and methodologies necessary to conduct probabilistic assessments.

In response to the recommendations of the SAP, OPP began a new initiative to develop and validate tools and methodologies to conduct probabilistic assessments to address terrestrial and aquatic risk. The purpose of this initiative is to strengthen the core elements of the ecological assessment process by developing and validating probabilistic assessment tools and methodologies. These methodologies are intended for use by OPP for evaluating effects of pesticides to terrestrial and aquatic species.

Ms. Keehner acknowledged the hard work and efforts of ECOFRAM, which have resulted in the completion of their draft aquatic and terrestrial reports. The purpose of this workshop is for the Peer Input Panel to review the Aquatic Workgroup's report. The Panel was specifically asked to consider these five questions:

* Is the report scientifically sound?
* Did the ECOFRAM working group fulfill its goals?
* What are the limitations to this approach?
* What areas need to be strengthened?
* How should this process be integrated in the overall risk management process?

**Terrell Barry, Ph.D., California Department of Pesticide Regulation, Environmental Monitoring and Pesticide Management Branch**

The report is scientifically sound, and the methods presented are standard and generally accepted. Chapter 2 gives a clear overview for those who need to understand the registration process. Chapter 4 is well organized, with much information, but confuses exposure concepts and risk characterization concepts. A set of standard selection guidelines is needed so that analyses can be performed in a uniform manner. The RADAR tool will lead to standardization of analyses. Natural stochasticity and parameter error are well covered, but model error should be discussed in more detail. The methods should be developed within the context of FIFRA and follow the outline in the Framework for Ecological Risk Assessment.

The main limitation to the report is too heavy a reliance on simulation modeling of exposure for tiers 3 and 4 without a clear indication that the models perform reliably in all expected scenarios, specifically the basin scale. Unresolved issues remain surrounding selection of simulation model input variables. More emphasis should be placed on the value of monitoring data and more time spent to develop monitoring methods. International harmonization is desirable but will take time.

The report could be strengthened through restructuring according to the U.S. EPA Ecological Risk Assessment Guidelines and other regulatory mandates and programs. Thus, a separate chapter on risk characterization should be added that addresses the joint probability curve concept and uncertainty. A statement identifying the intended audience of the report should be included. An unresolved issue is how regulators, the regulated community, and society can better understand which water bodies need to be protected and to what degree.

Initial mitigation measures can be proposed in tier 2 based on the joint probability curve. A well-characterized response profile coupled with a conservative exposure profile would enable evaluation of the need for risk mitigation.

**Peter Delorme, Ph.D., Health Canada, Pest Management Regulatory Agency, Environmental Assessment Division**

Canada's Pest Management Regulatory Agency is equivalent to EPA. Because of NAFTA, the report will have an impact in Canada. The report brings together a large amount of information. The tools and methods appear to be sound and are already in the literature. The tiers progress logically from simple through complex. The tier process and its tools and methods represent a reasonable first step toward probabilistic assessment. All models should be fully validated, however. Validation of exposure models, specifically, is achievable with resources and effort. Other components relating to ecosystem dynamics will be more difficult to validate. Methods to account for the effect of environmental fate factors on toxicity are included in the report, but species sensitivity is addressed only at tier 3. The report should account more fully for fate and effects of transformation products, effects of formulants, and formulation types. Events are happening in the field that would not be picked up in a risk assessment.

Lack of data is the biggest limitation to the probabilistic approach. We must be careful not to create a false sense of certainty by generating volumes of exposure data from unvalidated models. How the "one size fits all" farm pond scenario, for example, relates to actual field situations such as a multiple crop/single pesticide scenario is questionable. A range of water bodies needs to be addressed. Tier 1 scenarios need to be developed for other crops and regions.

Recommendations should be prioritized into major task groups and research needs. To educate risk managers, case studies should be developed and presented using typical levels of information--a case study on permethrin is a good example. The categorization of uncertainty should be presented earlier in the report. As each tool and method is introduced in the report, an accompanying statement should relate it to the rule to which it applies. Both exposure and effects workgroup members should advise on the kinds of water bodies to include in the scenarios.

The certainty level at tier 2 would support consideration of risk mitigation, in cases where the risk assessment indicates the product is close to acceptable. Risk assessment is very product-specific in many cases.

**Steve L. Foss, Washington State Department of Agriculture, Registration and Services Branch**

A State risk assessor, Mr. Foss has to make decisions on the "front line." The report contains many different models, values, and curves (e.g., GENEEC, risk quotients, estimated effect concentrations, level of concern values, and joint probability curves) about which risk assessors and managers need training. A trained risk assessor could then make decisions after a determination of risk at the tier 1 and 2 levels. Assessors need access to data and help in choosing the relevant data. For example, the "USDA-ARS Pesticide Properties Database" may serve as a useful resource of data. Guidance is needed on how to run models when there are missing data values. The report's presentation of the GENEEC model was clear. A person with minimal training should be able to run the GENEEC model by following the stepwise process and make proper decisions.

The biggest limitation, particularly in later tiers, is expense. We need modeling guidance for applications directed to water bodies--estuaries, lakes, and canals. The models should be expanded to include these sites with direct applications to water. The biggest challenge in regard to monitoring data is how to respond to pesticide detections and communicate risk. For example, what is the risk associated with the discovery of diazinon in urban streams by USGS in Washington State? Duration of exposure should be considered along with concentration in determining environmental significance.

An example was provided of prioritization of water bodies in Washington State. Although different agencies may have their own missions and objectives, there is a need for agreement on the risk assessment process in order to make consistent determinations. In Washington State, risk can be mitigated at several points in the registration process. When a registrant (applicant) requests an emergency exemption from registration (Section 18), the state has chosen to include a buffer as a mitigation option when the proposed use may pose a risk to endangered species and no risk assessment has been submitted. A list of mitigation measures from other States or EPA would be a helpful resource.

**Kathryn Gallagher, Ph.D., U.S. EPA, Office of Pesticide Programs**

The report presents a good basic framework for conducting probabilistic assessments. The suggestions given by the workgroup for general improvements of exposure assessment tests and tools are very useful, and the model descriptions are generally good. Some of the models need more validation, particularly GENEEC and EXAMS. Well-validated models are crucial. Variability in model input parameters should be considered, but was excluded. The expanded scale of exposure analyses described in the report could lead to overlooking negative impacts in sensitive, small-scale areas. Monitoring studies and field studies are minimized in the report but are important. The “highly conservative” model estimates do not always exceed actual environmental levels. Risk underestimation could result from the exclusion of degradates from modeling and monitoring, the lack of consideration of groundwater transport of chemicals into surface water bodies or irrigation water sources, the lack of consideration of diet and sediment as additional potential sources of exposure, lack of consideration of use of a pesticide on multiple crops in the same watershed, as well as the exclusion of consideration of the potential effects of exposure to multiple pesticides.

Strengths of the effects chapter include use of the dose-response curve, inclusion of Daphnia chronic and fish early-life stage tests in tier 1, and addition of higher tier tests. Tier 2 should require additional data on species sensitivity within and between taxa, variations at different life stages, and variations in effects with different formulations. Vertebrate chronic toxicity should be considered earlier than tier 3, possibly in tier 1. Sediment toxicity should be addressed in greater detail, as should semiaquatic species and indicators of sublethal effects (e.g., increased susceptibility to disease). A concern is that the outlined process may lead to endless refinements, with a concomitant consumption of resources. Discussion of a mechanism to remove chemicals from further consideration due to unacceptably high exposure or effects values should be included. Variability in exposure and effects should be addressed in the discussion of the joint probability curves. A number of suggestions for improving the draft report were given.

**John P. Giesy, Ph.D., Michigan State University, Department of Zoology**

Drafting the report was a momentous task. The tiered approach and the toolbox will accommodate improvements as the science grows and give flexibility at higher tiers. A closer tie to the ecological framework document would be helpful. The dilemma in dealing with pesticides is that we are working with altered ecosystems. Linking this document to the framework document by following the risk characterization process will help in the choice of appropriate scenarios and methodologies.

The probabilistic approach is not a perfect answer, but it is sound, more refined, and more realistic. It does not lock us into individual point estimates. Under this approach there are no absolutes. On the other hand, lines are blurred and there will be impediments to implementation from the regulatory perspective. The proposed framework will encourage collection of more information. Probability approaches are data intensive, particularly at monitoring and verification stages. Harmonization of base datasets among countries is needed for coherence. New chemicals with novel chemistries have few data. Population and community-level modeling do not provide enough accuracy. They are a useful framework, but in the regulatory context their parameters are too uncertain and they may never be useful in a quantitative manner. The variation in responses from these models is huge.

Mitigation should not be attempted until tier 3 is completed, because earlier tiers do not provide enough resolution or realistic estimates of exposure.

Improvements to the report are to integrate reciprocity analysis into the overall framework toward a toxicity curve/reciprocity relationship, and to refine species sensitivity.

Probabilistic risk assessment will continue to improve through incorporation of the latest information.

**Robert Graney, Ph.D., Bayer Corporation, Environmental Research**

Of most concern from industry's standpoint is the practicality of the framework and its concepts. The risk assessment process must be applied in a variety of case studies to determine its use under actual regulatory circumstances. Risk managers must be involved because they have to know whether the information will help in making decisions. Given the focus of the new approach on probabilistic assessment, it should be noted that many recommendations and components within the document are not directly related to probabilistic assessment. The recommendations were logical on the need for more fate and effects tests, but more guidance is needed on when to use such tests. Cost is an industry concern. Outdated studies should be identified and eliminated from the new process.

Ensuring consistency of risk estimations for compounds in tiers 3 and 4 will present a challenge to EPA management, given the flexibility of design. Models are not available yet for higher tiers. Three overall limitations are availability of adequate data, actually developing and maintaining the required exposure models, and training personnel. Including an executive summary in the report is critical, as is standardizing terms and concepts between aquatic and terrestrial risk assessment, reducing redundancy, and prioritizing recommendations.

Mitigation is an ongoing process that occurs naturally during all phases of risk assessment, depending on the product or the hazard. The minimum information necessary would come from PRZM/EXAMS analyses.

The risk manager's role in determining the need for additional data must be clarified. The need for additional data in tiers 3 and 4 is determined by unacceptable uncertainty, which must be better defined. Assessment endpoints must be defined; endpoints for all insecticides seem to vary from beginning to end of the report.

Suggestions for chapter 4 include redesigning the current fish study, rather than designing new studies; making some terminology consistent from tier to tier; and rewriting and clarifying the section on sediment toxicity tests. Mesocosms/microcosms provide a tremendous amount of information once they are designed, but they are very costly. Experience has been varied on whether improvements have resulted from mesocosms/microcosms.

**Dr. Jan B. H. J. Linders, RIVM-CSR, The Netherlands**

EPA is very directed to the U.S. situation; because the rest of world follows EPA, it is hoped that the agency will look beyond the United States in the future. The document reflects thorough work and is an excellent status report. Detailed guidance is somewhat overdone, and some chapters are redundant. Key items should appear in an executive summary. Processes should be described with an equal level of detail (e.g., the section on photolysis contains too much detail). Repetitive items, lists, and figures confuse both the reader and the user.

In reference to limitations for predicting risk, the main point is the degree of trust that managers have in the models' ability to characterize reality. Considering and proposing risk mitigation measures is timely at any point in the process. A company may propose mitigation in an early tier because of the time and money in running additional tests or calculations.

Chapter 2 provides an excellent overview of the tier process. However, analyses indicated for later tiers should not be postponed if data are available at early stages. The data available concerning ecotoxicological testing should be used to the maximum extent, which means full use of the joint probability curves because this is just a matter of handling existing data. Tier 2 does not seem sufficiently discriminatory. Another key point is international harmonization--the new probabilistic approach is unfamiliar in European regulatory processes. A criticism is that this report did not follow recommendations proposed several years ago regarding harmonization of data formats.

In reference to exposure assessment, a published RIVM report (#679101022) may be useful for determining environmental endpoints; it provides key parameters for each environmental study type. In the United States, field study data are regarded as difficult to interpret, but in Europe they are commonly used to resolve doubts in the registration of pesticides.

**Margaret Maizel, NCRI-Chesapeake Inc.**

The report was excellent, particularly the commitment to flexibility, the well-described procedural frameworks, and the emphasis on communication. Risk assessment managers and regulators need to communicate. The imbalance of information between the exposure and effects chapters should be addressed. As a GIS landscape analyst, Ms. Maizel agrees with the recommendations in chapters 3 and 4 for using landscape scenarios in model design, specifically as could be developed using a national-in-scope, local-scale, GIS-based integrated information system such as has been developed using the OSIRIS Information Utility (OIU).

The GIS analytical framework was first developed in 1987 for the National Resources Inventory. In 1992, in collaboration with USDA, the first groundwater vulnerability index for pesticide leaching below the roots was built on field-level data applied nationally. Data similar to those recommended by the ECOFRAM study group were added over the years from many agro-eco sources to describe arrays of ecological-environmental landscapes, using a hierarchical approach that is compatible with the proposed tier system. The OIU "looks from the nation to the neighborhood." The OIU was designed for modeling ecological systems in agricultural environments. Corn in Illinois was used, for example, as a model "resource management domain"--an area that should behave predictably around the world given certain management uses and landscape inputs.

This approach can be used to identify the universe of crop-specific landscapes that are relevant to models. It is possible to (a) incorporate real-world inputs into the models rather than fitting the landscape to the model; (b) look at natural landscapes, which need more comprehensive characterization and potentially to infer arrays of biological organisms potentially exposed within these domains; and (c) possibly address uncertainties in mesocosms and microcosms as being representative of agro-ecological systems..

It is suggested that probabilistic risk assessment address plume modeling more holistically; combinations of surface and groundwater effects, together with potentials for phenomena such as "mounding" effects; and combined interactions of effects of different pesticide uses in the same landscape unit.

**Miachel Rexrode, Ph.D., U.S. EPA, Office of Pesticide Programs**

Probability assessment tools are useful only if they reduce uncertainty. Momentum is building for probability assessment as a means of characterizing variability and uncertainty in fate, transport, exposure, and dose-response. In addition, the degree and probability of effects and exposure can be predicted. Under the current deterministic model, hazard potential is first determined, then risk assessors usually move directly to the fate and transport model

(PRZM/EXAMS). From here, the information is consolidated into a risk characterization. This approach does not address the number and distribution of vulnerable species or the degree of toxicity.

The draft report needs refinement to become an implementation tool for probabilistic assessment. The uncertainty of probabilistic models to be used in tiers 2 and 3 should be addressed by creating programs to validate the models. In addition, tiers 1 and 2 would be more useful with additional data. Data on most chemicals are now in the open literature. Extracting useful data from the literature is a challenge, however, as is determining how to use such data in a risk assessment. Very few data are available on new chemicals.

Data concerns include characterization of toxicity to benthic fish species and inadequacy of daphnids as a surrogate for all species. Probabilistic risk assessment will benefit by including sediment testing on representative invertebrates, amphibian information, dose-response slopes, and accuracy of the current toxicity test.

Tiers 1 and 2 should be combined with a probabilistic assessment using an adequate effects database and validated exposure models, and with uncertainty defined at each level. Those who use these models need definitions of uncertainty for conducting risk assessment.

**William van der Schalie, Ph.D., U.S. EPA, Office of Research and Development**

ECOFRAM has done a good job in expanding available tools for pesticide risk assessment. It did not consider tools to use at increasing levels for indirect effects, however, perhaps on the basis that extrapolation would be performed for indirect effects. The report should mention the suitability of ecosystem models for this application.

Terms used in the report should relate to the 1998 ecorisk guidelines, and the document should be structured according to the ecorisk process. Some of the tools for analyzing effects need validation, and they may require additional data. The methods need to be standardized, and should be specific enough to allow different risk assessors to estimate similar values of risk. Tiers 3 and 4 are flexible, and their lack of standardization may not be a disadvantage. The methods are explained clearly, although some may include too much detail. In addition, assessments in tiers 3 and 4 may vary considerably depending on the chemical being analyzed. Thus the rationale for chosen scenarios should be explained, along with any assumptions or uncertainties. Development and validation efforts should be better defined; approaches in tiers 1 and 2 need retrospective evaluation through application of case studies and existing data.

The report would benefit from discussion on assessment endpoints, appropriateness of LOCs, geographical specificity, mixture toxicity, and utility of ecosystem models. A separate section on risk characterization should be added. New approaches should be validated using existing data.

Risk mitigation might be addressed at tier 2, with the help of joint probability curves. All involved need to reach consensus on assessment endpoints and acceptable uncertainty.

## PANEL DISCUSSION

Panel member discussion on these topics is summarized in this section. ix primary areas of concern were identified from the peer reviews, as listed below.

1. Assessment Endpoints: Ideal versus Practical
2. Aquatic Risk Assessment Process
3. Spatial Scale
4. Temporal Scale of Effects and Exposure
5. Risk Characterization: Joint Probability Curve or Other Approaches
6. Validation

### 1. Assessment Endpoints: Ideal versus Practical

Preestablished endpoints clarify what is being protected and thus what is being assessed. Endpoints to include in risk assessments need to be defined. The aquatic report workgroup recommended population and community levels as endpoints for the early tiers and did not think endpoints should be refined because these would be different for each chemical assessed. The report stated that measured endpoints become more complex at the higher tiers.

*Generic or Specific Endpoints for Tiers 1 and 2?* Discussion addressed whether generic endpoints are satisfactory for tiers 1 and 2. Some thought that generic endpoints help everybody involved understand what is going on. Generic endpoints should be applicable to any ecosystem and redefined according to what issues need resolution as one moves through the tiers. Exchange and interchange between tiers are common. Tier 1 should be conservative and protective. But even in Tier 1, it was noted that identifying more specific endpoints is worthwhile if details are available.

Others discussants thought there are too many unknowns in the first two tiers, the tools are limited, and the database should be increased. Defining specific endpoints early in the assessment gives better structure throughout the process. It might also better identify problem compounds. A case was cited of a new problem pesticide on the market that would have been cleared at an early stage under the proposed system but that could have been caught through sediment testing. Perhaps potential problem chemicals should be identified in advance so they are not missed in tier 1. It was suggested that tier 1 endpoints address the sustainable part of the population needed to maintain a given ecoquality, for example, diversity of a fish population. The indirect effects on quality of fish population would be a consideration.

It was argued that only at tiers 3 or 4 can such questions be answered, because tiers 1 and 2 include only one or two fish species. If one has a concern, one moves to the next level rather than refining endpoints at lower tiers. At tier 3, one might look at, for example, functionality of ecosystem or diverse fish population groups.

Experience has shown that if a risk assessment seems to be showing theoretical impact at tier 2 through the use of generic endpoints, a risk manager might decide to mitigate. With limited data at that stage, however, the mitigation might be insufficient to deal with actual risks. If there were earlier specific data, the risk management decision might be different. It should be noted that registrants often avoid going as far as tier 3 or 4 because of expense. From the industrial standpoint, if spending resources to obtain data will lead to greater predictability earlier, then such expenditure may be worthwhile.

*Who Should Establish Assessment Endpoints?* A successful strategy has been for the scientist to suggest possible endpoints to the risk manager, thus initiating a dialogue. Scientists can help risk managers determine what might be a problem compound and can explain the biology and ecology if necessary. Several EPA activities are seeking to identify common important endpoints that might be useful to this effort.

*Structure vs. Function Endpoints.* An area of disagreement is whether endpoints should relate to community structure or function. Both the current risk assessment system and the proposed one lead to structural endpoints, which are easier to define. Lower tiers should be limited to structural endpoints--this is easier from the standpoint of the risk assessor and the regulator. Implementing a measurement endpoint that is coupled with a functional assessment endpoint is difficult. Functional endpoints are in general less sensitive, but should be maintained. This is not well articulated in the document.

*Standardization of Endpoints.* Regulators in localities across the United States and abroad need endpoints applicable to their local aquatic environments and local mandates. These endpoints need to be identified. Panelists suggested the following endpoints:

- Headwater systems
- Aquatic community structures
- Areas likely to have the most density of aquatic use next to them
- Special places/special habitats--areas with high ecological value
- Protection of fish
- Recovery – temporal scale

## 2. Aquatic Risk Assessment Process

*Content and Value of Tier 1.* The tier 1 goal is intended to be more protective than predictive. GENEEC, the tool used in tier 1, is inexpensive and quick to run. The majority of chemicals assessed are passed on to tier 2. Incorporating tier 2 effects analysis into tier 1 is favored by some discussants, because the additional effort to assess tier 2 effects is not substantial. Current databases could serve as a source of information, if they meet accountability and quality criteria. Some tests are inexpensive. The tier 2 tool, PRZM/EXAMS, is somewhat more intensive and expensive. Although actual run time is 40 minutes, full assessment takes 10-

14 days because of the need to reassess and perform additional analyses. MUSCRAT, another tool, is highly automated and involved.

It should be noted that ECOFRAM never intended to limit a risk assessment if higher level data are available--appropriate studies may be brought in at any time even though specified for higher tiers. However, any risk assessor using the information would have to be able to reach the same conclusions.

*Who Does Risk Assessments?* Should industry registrants submit risk assessments to EPA for approval or should the agency conduct them? This was not specifically addressed in the report except to state that risk assessments should be "submissible." The agency opinion is that the registrant community could conduct the initial risk assessment using standardized EPA-defined models and assumptions. If a compound will move to higher tiers, the registrant should obtain the data before submitting the risk assessment to EPA. This is more efficient. The registrant should address any needs before submission to EPA. It was noted that it would be naive to think that a registrant would complete a risk assessment that would be unfavorable to its product.

ECOFRAM's main goal is to provide a standard, clear, efficient system that enables the industry registrant to assume the burden of risk assessment while accomplishing EPA goals. Theoretically, enough detail should be provided such that a risk assessment from industry would be the same as one from the agency. As long as guidance is provided, computer modeling should give everybody the same results, whether industry or EPA. A centralized website would provide access to all. Pesticide risk assessment has in the past involved considerable "personal art." Even within the proposed approach, the toolbox concept presented in tier 3 incorporates flexibility, which was intended as an advantage. Such flexibility, however, could lead to arguable results. Standardization can only go so far, particularly at tiers 3 and 4, which are very expensive. Registrants need communication with the agency at these tiers.

In Canada there is no tiered system. The Pest Management Regulatory Agency requires a package of studies from companies.

*Building Monitoring Into Risk Assessment.* Discussants provided several viewpoints on the use of monitoring. The issue is the value of model estimates over limited monitoring data. A positive perspective is that monitoring warrants more attention from ECOFRAM. Use of monitoring data depends on what is being protected, however. Monitoring provides a good opportunity to obtain data on a large river system, for example. In some areas, monitoring data are more useful than data generated by models. Investing time and effort in monitoring might be more valuable than in validating models. Other agencies are generating monitoring data, and ECOFRAM does not need to design new programs. It makes sense to look at existing data to ascertain where problems are.

The other viewpoint is that monitoring gives only a general idea of concentrations because of the probability that peak concentration is not captured in infrequent tests. To be effective, monitoring must be compound-specific with focus on a particular issue. A recommendation was to include sediment testing in monitoring efforts, to better reveal concentrations of certain chemicals. Risk assessors have found it extremely difficult to make accurate exposure assessments from even a good monitoring dataset. Monitoring data are lacking for smaller watersheds. Empirical models were suggested for small watersheds--a cross between a stochastic and a regression model, coupled with mechanistic models. It was pointed out that headwater data can be captured in a small study and might be useful for calibrating models.

Appropriate uses of monitoring data must be determined, and limits must be set. General monitoring studies should be differentiated from specifically designed field tests.

### 3. Spatial Scale

An important ECOFRAM recommendation for the higher tiers of probabilistic risk assessment is to develop a tool to account for variability in the landscape--exposure scenarios, different sizes of watershed, weather variations, and so on. GIS information is a resource that defines a typical landscape to enable extrapolations to other parts of the country. GIS can be used to mitigate certain use patterns, identify sensitive spots, and ascertain problem patterns. Surface water management can benefit from GIS as well. One discussant viewpoint was that GIS data are helpful to benchmark and extrapolate but should be used along with monitoring data.

Capturing local variation within a spatial scale is a concern, for example, soil type within a field or hydrological properties of a lake, such as spatial variation of runoff, which does not mix into the lake water instantly. Affected small-scale areas should not be lost. The point was made that minor crops are often very intensively pesticide-treated and are associated with heavy concentration peaks.

### 4. Temporal Scale of Effects and Exposure

The temporal scales in the exposure and effects chapters of the draft report should match. The use of peaks versus time-weighted averages was not resolved. At issue was whether protocols should be expanded from typical LC50s to time-to-event analysis. ECOFRAM proposed only minor modifications. Unresolved factors include multiple exposures, pulsed short-duration exposures, and short-term versus long-term exposure. It was noted that time-weighted averages give better values for chronic exposure. One discussant stated that new protocols have the potential to determine the probability of mortality. Another questioned whether time-to-event analysis will lead to better conclusions. Its use makes sense to academicians and scientists, but will require considerable retraining. Given the highly variable data, considerable resources might be consumed to change the protocols.

### 5. Risk Characterization: Joint Probability Curve or Other Approaches

Joint probability curves express risk characterization by showing the relationship between magnitude of effect and probability of occurrence. The curves are tools that provide a transparent estimate of risk that can help any side make a decision. Because the curves provide much more information and realism, decisions based on the curves must include judgment and good will. Risk managers need guidance on how to choose appropriate curves, determine when an assessment is finished, and recognize a worrisome area on the curve. The vertical axis of the curve is harder to communicate than the horizontal. One must define in advance what one wants the axis to mean. Case studies are highly recommended as a means to illustrate use of the curves to managers. If the assessment endpoints are established initially, the number of curves generated is reduced.

*Number of Species.* Discussants addressed the number of species needed to create a curve--seven to nine species was deemed a reasonable number. When only one or two species are available, which is common, uncertainty must be addressed. One solution is to apply an uncertainty factor--a confidence bound--that covers the variability in the data. This method has been used in terrestrial risk assessment but not in aquatics. It was pointed out that aggregating the data gives a misleading estimate of probability of effect in a particular target group. Also, curves for sensitive invertebrates should not be mixed with those for tolerant fish. It was suggested that more species be added in tier 1, particularly for a compound with major widespread use. Risk assessors and managers are more willing to accept uncertainty with a minor compound.

*Mesocosms.* Several discussants mentioned the use of mesocosm studies, which include information on a number of species at once and could be helpful in later tiers, where the daphnid as a generic surrogate may not best represent other species. Mesocosm studies are expensive and high risk, although new protocols are reportedly more manipulable and have fewer problems. Industry is more willing to sponsor mesocosms at an early tier, when they are simple and less costly and may lead to a better, faster decision. In Europe, when the first tier for aquatic organisms is exceeded, most agencies conduct mesocosm studies. The section on mesocosm studies in the draft report should be clarified.

## 6. Validation

The consensus in the reviewer comments was that model tools and process need validation. The value of validation should not be overestimated, however. No matter how well a model is validated, it still is a simplification of reality even though it contains all the scientific information currently available on the behavior of pesticides in the environment. The benefit of models such as PRZM/EXAMS is that all possible scientific information goes into a decision and individual judgment is reduced. However, risk managers want assurance that the models accurately reflect the environment. In industry, individual scientists are skeptical, because they have to defend the models to the risk managers. The models are intended to save time and money and thus should be accurate before they are relied upon. Normal development of a model includes validation, but budgets sometimes run out before validation is completed.

Datasets and techniques are available for use in validation. Studies being conducted by various government bodies are yielding data that could be forged together. Field study data have been found hard to use for model validation, however. Comparing field study results with results from a numerical model is difficult. Furthermore, the components of a field study needed for validation of a model often are dropped for monetary reasons.

Risk assessors and managers are under constant pressure to regulate new sites and often are forced to use model tools--validated or not--for purposes and locales other than those they were designed for. When used for new and different scenarios, the models may not be as conservative as expected. The scenario set up initially defines the conservativeness, and the scenario must be accurate. A common framework would be ideal, into which scenarios are plugged. First, scenarios should be identified and then modeling schema improved. The level of risk estimated depends on the site/scenario picked; thus, the scenario is a critical component of the risk assessment process. Landscape analysis should help with identifying scenarios.

Confidence in the process can be bolstered through case studies showing how models did or did not work. Models do not work equally in every part of the country, because of local variations. Given the many components associated with the models, using a task force or larger work group approach to validation might increase confidence. The group perspective is more convincing. ECOFRAM should develop clear instructions on use of the models and guidelines on using the process, so that it is reproducible. Most instructions come from the modelers.

The point was made that registrants, not EPA, should bear the burden of demonstrating compound safety and thus assuring confidence. Experience has shown that risk managers usually express lack of confidence only when scientists think there is a problem--then the managers want validation. Tiers 3 and 4 require more confidence than tiers 1 and 2. One participant doubted the usefulness of a distinction between risk assessors and risk managers.

## PRIORITIES

Panel members were asked to list priorities to assist EPA as it moves forward to implement the recommendations in the document.

**Delorme**

- Look at assessment first, which will drive amount of data needed in tier 1
- Increase data at lower tiers
- Conduct case studies to see how everything fits together
- Conduct validation
- Educate risk managers

**Foss**

- Valid models—run scenario for treatment to lakes/rivers
- Educate risk managers
- Make data available on internet

**Gallagher**

- Conduct case studies that include effective different approaches--compare field-scale effect with basin-level effect

- Include targeted exposure for model validation–do additional field experiments such as at mine sites to look at runoff
- Require additional effects data at lower tiers–more species or life stages
- Pay attention to degradates and other compounds that have a common mode of action

**Giesy**

- Endorse monitoring data, but consider using other terms for monthly and semiannual monitoring such as "field data" or "field-scale information"
- Conduct case studies--targeted field measurements, linked to events

**Graney**

- Define assessment endpoints: population, community, ecosystems
- Apply ECOFRAM approaches in case studies, then prioritize
- Focus on tier 2 models--MUSCRAT must be upgraded; develop standardized and transparent scenarios
- Establish acceptable uncertainty--identify triggers in each tier indicating acceptable uncertainty

**Linders**

- Update all tools with additional scenarios and make tools easier for non-EPA users
- Use available data to the extent possible (e.g., new developments in time-to-event analysis)
- Develop necessary probabilistic tools--the ones mentioned are not really operational
- Provide education on use of the new instruments

**Maizel**

- Incorporate landscape units in risk assessment--to contribute information on (a) habitat type, (b) waterbody type, (c) organism arrays that can help with joint probability curves, and (d) selecting field test areas to help in model validation

**Rexrode**

- Develop a complete database--limited data bring limited certainty
- Further refine processes
- Conduct case studies to provide confidence

**Schalie**

- Develop case studies first rather than focus on specific recommendations
- Define endpoints now
- Develop a website now to improve communications

- Decide on the basis of case studies what recommendations to accept

- Triage individual recommendations for short term versus long term

## PUBLIC COMMENT

**David Esterly, DuPont**

ECOFRAM has done a commendable job in sorting out all the variables. Having developed a scope of work, ECOFRAM now has an obligation to answer the questions it posed in the document. A statement of procedures should be developed for risk assessments. At DuPont, the risk managers need numbers. The risk assessment process would be less confrontational and more cooperative with a common set of tools.

Environmental toxicology concentration numbers must be based on description of environment. Description of the environment that we are trying to protect is significantly lacking. If the farm pond is to be used as a representative environment for registration, the associated arithmetic should not be based on the flowing stream environment.

A joint probability distribution provides a value for frequency of failure or of an event occurring. This number must be understood in the context of either number of acres or application. Acres in production must be considered for runoff, and application must be considered for drift. In reference to acceptable failure rate for an application, one must ascertain whether the cause was human error or uncontrollable environmental consequence.

Regarding model validation, PRZM runoff is based on empirical equations generated over many years. One does not need to revalidate PRZM because the original datasets show how the validation was done. Another side of the validation issue has been missed: comparative modeling. Comparison of a problem product with an acceptable product enables identification of the seriousness of the situation.

**Richard Lee, U.S. EPA, Environmental Fate and Effects Division**

The objectives of ECOFRAM are twofold: (1) to consider probability (uncertainty) in statistical inferences and (2) refinement of test procedures as the tiers progress with consideration of real-world scenarios (e.g., population abundance, species diversity, trophic interaction). However, the following setbacks are unavoidable if excessive test procedure refinements are attempted:

- It will become a specific (local) case with greatly increased uncertainty (with accurate test results, but reflection of only the local scenario).
- It will result in a single data point on temporary and spatial scales.
- Validation, time, and expense will become issues.
- It will result in voluminous, inconclusive information with a significant gray area (uncertainty) based on our experience with mesocosm studies.
- To be idealistic (but not practical), we have to test all formulated products instead of

benchmark policy.

- The additive or even synergistic effects of pesticides in the field may become an issue.
- For population/ecosystem modeling and field testing (proposed for advanced tier testing), no refined and complex model is currently available, and availability of necessary data is also questionable. In the past, a problem for the field testings has been with insufficient replications.

To reduce uncertainty (with a probabilistic concept) we should consider basic issues such as improvement of data quality. The followings is a list of do's and don'ts:

- Utilize all available confidence intervals (CIs) for stochastic inference. Transfer information from CIs to standard normal distribution for calculation of probabilities.
- Since we have CIs for every $LC_{50}$ or $LD_{50}$ and $EC_{50}$, why can't we express RQ with the range?
- We have a tool for an aquatic EEC Monte Carlo simulation based on soil and hydrological variables. Why not add the chemical factor such as $K_{t1/2}$ (or the other factors via sensitivity analysis)?
- For statistical analysis of hypothesis testing (such as T-test, ANOVA , or ANCOVA), consider:

    - Power of test (degree or probability of confidence)

    - Alternative hypothesis testing recommended by Dr. Stunker to eliminated the grey area for which we are not responsible or simply increase sample size

    - Start to conduct a normality test and homoscedasticity test.
- For probit analysis or regression analysis (for LCx determination) consider:

    - Goodness of fit ($X^2$ )- test) and positive regression response (T-test)

    - Discontinue use of moving average method or binomial method for $LC_{50}$ determination.

    - Do not use $LC_{10}$ or $LC_5$ specially with flat slope, because it will greatly increase uncertainty.
- The bottom line for improving the data quality and test accuracy is to increase a data sample size.

**Daniel Rieder, U.S. EPA, Environmental Fate and Effects Division**

ECOFRAM should differentiate between the new data requirements for probabilistic risk assessment versus the data needed to expand the scope of the risk assessment process. In addition, ECOFRAM should focus on implementing probability risk assessment processes rather than rearranging tools in the toolbox. Refinements should be constructive.

## APPENDIX A
## AQUATIC PEER INPUT PANEL MEMBERS

Dr. Terri Barry
California Department of Pesticide Regulation
Environmental Monitoring and Pesticide Management Branch

Dr. Harold L. Bergman
University of Wyoming
Zoology & Physiology

Dr. Peter Delorme
Environmental Assessment Division
Pest Management Regulatory Agency

Mr. Steve L. Foss
Washington State Department of Agriculture
Registration & Services Branch

Dr. Kathryn Gallagher
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Dr. John P. Giesy
Michigan State University
Department of Zoology

Dr. Robert Graney
Bayer Corporation
Environmental Research

Dr. Jan B.H.J. Linders
RIVM-CSR
The Netherlands

Ms. Margaret Maizel
NCRI-Chesapeake Inc.

Mr. Miachel Rexrode
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Dr. William van der Schalie
U.S. Environmental Protection Agency
Office of Research and Development
National Center for Environmental Assessment

**APPENDIX B**
**AQUATIC WORKGROUP MEMBERS**

***Exposure***
James Baker
Iowa State University
Department of Agricultural and Biosystems Engineering

Lawrence Burns
U.S. Environmental Protection Agency
Office of Research and Development
NERL

David Farrar
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Paul Hendley
Zeneca Ag. Products

Alan Hosmer
Novartis, Ecological Toxicology

David Jones
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Walton Low
U.S. Geological Survey

Mark Russell
DuPont Ag. Products

Mari Stavanja
Florida Department of Agriculture and Consumer Services
Bureau of Pesticides

Martin Williams
Waterborne Environmental

James Wolf
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

***Effects***
Lawrence Barnthouse
LWB Environmental Services

Jeffrey Giddings
Springborn Laboratories

A. Tilghman Hall
Bayer Corporation
Agriculture Division

Michael McKee
Monsanto Company
Agriculture

Michael Newman
College of William and Mary
Virginia Institute of Marine Science

Kevin Reinert
Rohm and Haas Company
Environmental Toxicology and Ecological Risk Assessment

Robert Sebastien
Environment Canada
Chemical Evaluation Division
Commercial Chemicals Evaluation Branch

Keith Solomon
University of Guelph
Centre for Toxicology

Ann Stavola
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Leslie Touart
U.S. Environmental Protection Agency
Office of Pesticide Programs
Environmental Fate and Effects Division

Randy Wentsel
U.S. Environmental Protection Agency
Office of Research and Development

## APPENDIX C
## AQUATIC PEER INPUT PANEL COMMENTS